Course Homework (NMAI059)

Winter Term 2009-10

Petr Koupý, kouppe7am, petr.koupy@gmail.com

Seed = 14061987

Note: This document assumes that reader has already reviewed the source code written in R and is familiar with its structure. Source code is well commented and explains most of the reasoning. In short, source code provides single simulation function, which can be extensively parameterized. However, the only compulsory parameter is the identification number of the task to solve (task number corresponds to the order of the task in the homework assignment). Source code contains two simulations (one for each station), each consisting of definition of data structures and simulation loop. Simulation of Station Two (Luková) is dependent on results from simulation of Station One (Zahrádky). According to selected task, statistics and graphs are calculated after the simulation.

1.

a)

```
> chopok(task = 1)
Station One arrival count: 2087
Station Two arrival count: 1558
P[passenger at Station Two is from Station One] = 0.6662388
```

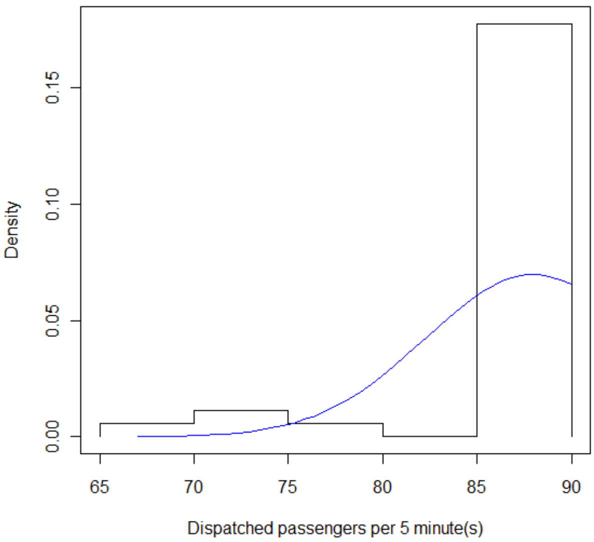
Since all proper (in time) arrivals of passengers at both stations must be served (even after the gates are closed), number of arrived passengers is equal to the number of dispatched passengers. We can see that the percentage of passengers at Station Two, who are transferred from Station One, implies that approximately half of passengers dispatched at Station One decided not to transfer to Station Two.

b)

```
> chopok(task = 2)
P[full load at Station One] = 0.937037
mean[load at Station One] = 3.864815
var[load at Station One] = 0.3174981

        1-sample proportions test without continuity correction

data:  oFullLoadCount out of length(oDepartureLoads), null probability
oFullLoadCount/length(oDepartureLoads)
X-squared = 0, df = 1, p-value = 1
alternative hypothesis: true p is not equal to 0.937037
95 percent confidence interval:
 0.9133037 0.9545963
sample estimates:
       p
0.937037
```

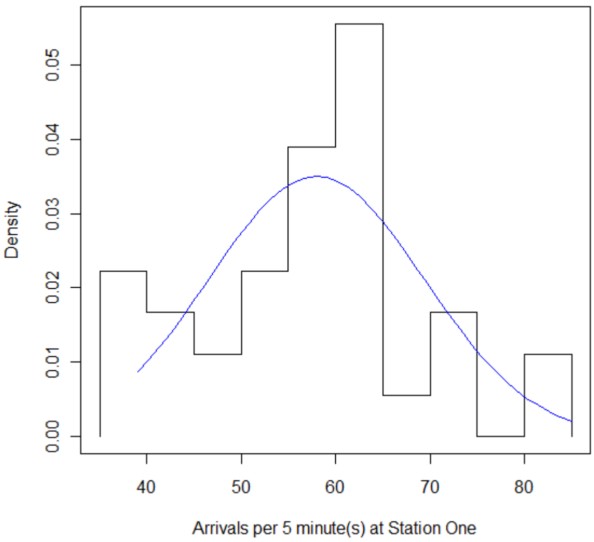
Provided statistics suggest that Station One is very heavily utilized almost all the time. When looking at raw data, non-full loads are very rare, usually appearing at the very beginning or when emptying the queue after the gate is closed.

c)
```
> chopok(task = 3)

        Shapiro-Wilk normality test

data:  dispatchSnapshots
W = 0.4061, p-value = 6.363e-11
```



Since P-value is very low comparing to usual 5% confidence, null hypothesis can be rejected based on the high significance of the result. It is apparent from the graph that measured data corresponds to normal density function poorly, although the peaks are aligned.

d)
```
> chopok(task = 4)

        Shapiro-Wilk normality test

data:  oArrivalSnapshots1
W = 0.9603, p-value = 0.2189
```



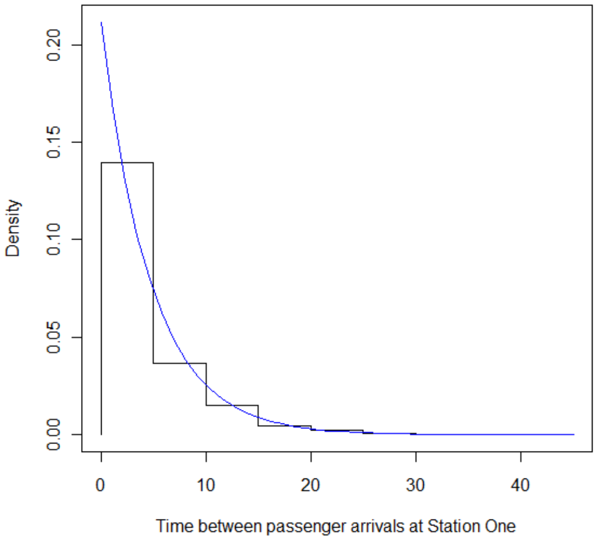
This time, the P-value is way above the confidence level of 5%. Thus, the null hypothesis cannot be rejected, assuming the data are normally distributed. Graph shows visible correspondence between measured data and normal density function.

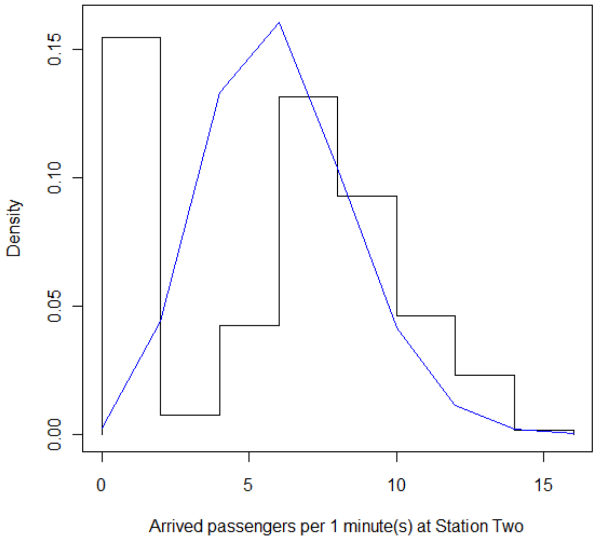
e)
```
> chopok(task = 5)
mean[waiting time at Station One] = 263.1049
mean[waiting time at Station Two] = 2493.068
```

Results suggest that mean of the waiting time at Station One is approximately 4 minutes, which is acceptable. Situation at Station Two is considerably worse - mean of waiting time is approximately 42 minutes. Considering results from a), it is apparent that Station Two must serve only quarter less passengers than Station One but with half capacity per car, which is clearly insufficient.

f)
```
> chopok(task = 61)
lambda = 0.2117092 (arrived passengers per second at Station One)
> chopok(task = 62)
lambda = 6.015444 (arrived passengers per 1 minute(s) at Station Two)
```

Poisson distribution applied to generating arrival count should produce the same results as exponential distribution applied to generating time between arrivals, assuming the



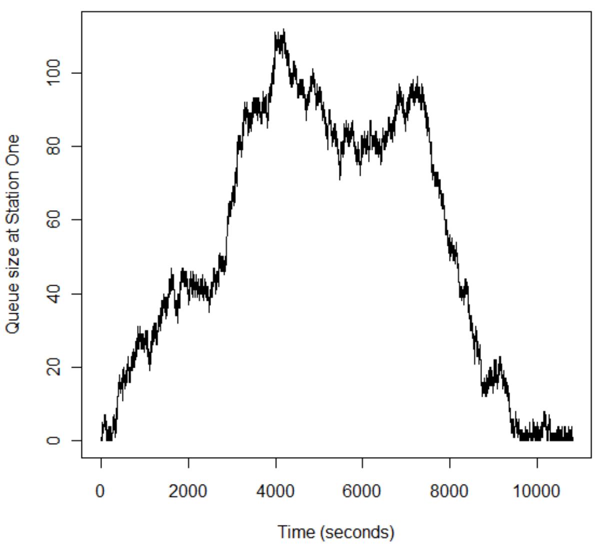


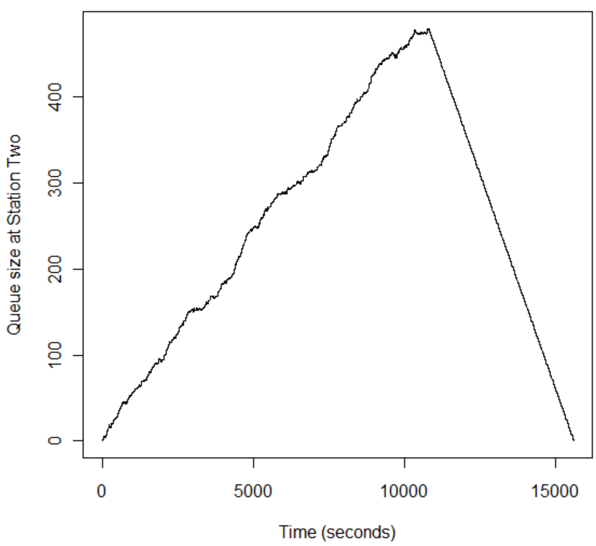
same lambda rate parameter. As we know, arrivals at Station One are produced by Poisson random value with lambda equal to 0.24, 0.2 and 0.16 respectively to corresponding intervals. Consequently, measured times between arrivals should correspond to exponential density function with lambda equal approximately 0.2 (see first measured lambda and corresponding graph above). Same reasoning can be inversely applied for Station Two, where time between arrivals (from environment) are generated by exponential distribution with lambda equal to 0.05. However, situation at Station Two is influenced by the transfers from Station One. Since we know from b) that cars from Station One are fully loaded most of the time, decisions of transferred passengers are distributed binomially (4 alternatively distributed trials every 20 seconds). Because binomial distribution with such parameters cannot be sufficiently approximated by Poisson distribution, resulting data does not match as well as in the first case.

g)

```
> chopok(task = 71)
> chopok(task = 72)
```

Similarly as in e), results shows that Station Two is not dimensioned sufficiently – queue size grows all the time until the gates are closed. Graph for Station One shows that capacity of the station is sufficient for middle interval (10:00 – 11:00), where passengers are dispatched at the same rate as new passengers arrive.





2.

a)

```
> chopok(task = 8, oCarInterval = 20, oCarCapacity = 4)
Station One
Car interval = 20
Car capacity = 4
Snapshot interval = 10 minutes
Percentages of arrivals over maximum dispatch rate of the station:
105 109 108 102 120 122 116 89 88 98 106 107 77 78 81 88 80 68

> chopok(task = 8, oCarInterval = 19, oCarCapacity = 4)
Station One
Car interval = 19
Car capacity = 4
Snapshot interval = 10 minutes
Percentages of arrivals over maximum dispatch rate of the station:
100 104 102 97 114 116 110 85 83 93 101 101 73 74 77 83 76 64

> chopok(task = 8, oCarInterval = 18, oCarCapacity = 4)
Station One
Car interval = 18
Car capacity = 4
Snapshot interval = 10 minutes
Percentages of arrivals over maximum dispatch rate of the station:
94 98 97 92 108 110 104 80 79 88 95 96 69 70 73 79 72 61

> chopok(task = 8, oCarInterval = 15, oCarCapacity = 4)
Station One
Car interval = 15
Car capacity = 4
Snapshot interval = 10 minutes
Percentages of arrivals over maximum dispatch rate of the station:
79 82 81 77 90 91 87 67 66 73 79 80 57 59 61 66 60 51

> chopok(task = 8, oCarInterval = 20, oCarCapacity = 5)
Station One
Car interval = 20
Car capacity = 5
Snapshot interval = 10 minutes
Percentages of arrivals over maximum dispatch rate of the station:
84 87 86 82 96 97 93 71 70 78 85 85 61 63 65 70 64 54
```

Measured data suggest that in order to satisfy requirements, it is necessary to shorten dispatch interval from 20 to 18 seconds. Both boundary cases (15-second interval or 5 seats) satisfy generated arrivals without creating the queue (at least not at 10 minutes resolution of data).

b)

```
> chopok(task = 9)
Station One arrival count: 2087
```

As was said in 1.a), number of arrived passengers is equal to the number of dispatched passengers. Theoretical expected value can be computed directly from provided lambda values, because Poisson distribution has expected value equal to lambda.

$$EX = 1.2 \cdot \left(\frac{3600}{5}\right) + 1 \cdot \left(\frac{3600}{5}\right) + 0.8 \cdot \left(\frac{3600}{5}\right) = 864 + 720 + 576 = 2160$$